

Original citation:

Grosskinsky, Stefan, Redig, Frank and Vafayi, Kiamars. (2013) Dynamics of condensation in the symmetric inclusion process. Electronic Journal of Probability, Volume 18 . Article number 66. ISSN 1083-6489

Permanent WRAP url:

<http://wrap.warwick.ac.uk/55455>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 3.0 Unported (CC BY 3.0) license and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/3.0/>

A note on versions:

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick**publications**wrap

highlight your research

<http://go.warwick.ac.uk/lib-publications>

Dynamics of condensation in the symmetric inclusion process

Stefan Grosskinsky* Frank Redig† Kiamars Vafayi‡

Abstract

The inclusion process is a stochastic lattice gas, which is a natural bosonic counterpart of the well-studied exclusion process and has strong connections to models of heat conduction and applications in population genetics. Like the zero-range process, due to attractive interaction between the particles, the inclusion process can exhibit a condensation transition. In this paper we present first rigorous results on the dynamics of the condensate formation for this class of models. We study the symmetric inclusion process on a finite set S with total number of particles N in the regime of strong interaction, i.e. with independent diffusion rate $m = m_N \rightarrow 0$. For the case $Nm_N \rightarrow \infty$ we show that on the time scale $1/m_N$ condensates emerge from general homogeneous initial conditions, and we precisely characterize their limiting dynamics. In the simplest case of two sites or a fully connected underlying random walk kernel, there is a single condensate hopping over S as a continuous-time random walk. In the non fully connected case several condensates can coexist and exchange mass via intermediate sites in an interesting coarsening process, which consists of a mixture of a diffusive motion and a jump process, until a single condensate is formed. Our result is based on a general two-scale form of the generator, with a fast-scale neutral Wright-Fisher diffusion and a slow-scale deterministic motion. The motion of the condensates is described in terms of the generator of the deterministic motion and the harmonic projection corresponding to the absorbing states of the Wright-Fisher diffusion.

Keywords: inclusion process ; condensation ; coarsening dynamics ; Wright-Fisher diffusion.

AMS MSC 2010: 60K35 ; 82C22 ; 82C26.

Submitted to EJP on April 4, 2013, final version accepted on June 22, 2013.

Supersedes arXiv:1210.3827.

1 Introduction

The inclusion process, introduced in [1] as the dual of a model of heat conduction, and further developed in [2], is a natural counterpart of the well-known and extensively

*Mathematics Institute, University of Warwick, UK.

E-mail: S.W.Grosskinsky@warwick.ac.uk

†Delft Institute of Applied Mathematics, Technische Universiteit Delft, Nederland.

E-mail: F.H.J.Redig@tudelft.nl

‡Department of Mathematics and Computer Sciences, Technische Universiteit Eindhoven, Nederland.

E-mail: k.vafayi@tue.nl

studied exclusion process. The inclusion process on a finite set S combines random walk jumps at rate $m p(i, j)/2$ with inclusion jumps, where each particle at site i is attracted by each particle at j independently with rate $p(i, j)$. Here $p(i, j), i, j \in S$ are the rates of a symmetric irreducible continuous-time random walk on S . The inclusion jumps introduce a form of attractive interaction between the particles.

When m tends to zero and simultaneously many particles are in the system, due to the attractive interaction between the particles, large piles of particles will be formed at individual sites. In fact we proved in [3] condensation for the stationary measure, in a limit where both $m \rightarrow 0$ and the number of particles $N \rightarrow \infty$. In this paper we prove dynamical results for the condensation phenomenon. More precisely, we study how clusters arise from an arbitrary initial condition, and how they move and merge into a single condensate, which then jumps over the finite lattice as a random walk. We therefore make $m = m_N$ dependent on $N \in \mathbb{N}$ such that $m_N \rightarrow 0$ as $N \rightarrow \infty$, which implies that diffusion will get slower and the attractive inclusion interaction will create condensates.

This model can be interpreted as a multi-allele version of the Moran model [4], describing the evolutionary competition of several species in a fixed size population. The inclusion part describes reproduction and death and m_N plays the role of an additional mutation rate, which is typically very small on the reproduction time scale (see e.g. [5] and references therein). It is also important as a dual process to models of energy or momentum transport [1, 2, 6, 7].

Related recent theoretical results include explosive condensation in a totally asymmetric model [8] which exhibits a slinky motion of the condensate also observed recently in [9] for non-Markovian zero-range dynamics. Results on the equilibration dynamics in zero-range processes with decreasing rates are currently under investigation [10]. In order to have a well defined limit dynamics we require symmetry of the $p(i, j)$, but our proof is not based on potential theoretic methods as used in [11, 12] for zero-range processes. Due to the system-size dependence of the diffusion rate we can apply more general multi-scale methods (see e.g. [13]). In particular, we use a two-scale decomposition of the generator and construct the generator of the limit process similarly to results in [14].

The decomposition consists of a Wright-Fisher type diffusion part which runs at “infinite speed” (in the limit $N \rightarrow \infty$), and a $p(i, j)$ -dependent drift part. The limiting motion is then described by the harmonic projection of the drift part on the absorbing set of the diffusion. In the simplest case of two sites, or similarly, in the fully connected case where all $p(i, j)$ are strictly positive, the absorbing set of the diffusion is the set of corner points of the simplex $E = \{x \in [0, 1]^S : \sum_{i \in S} x_i = 1\}$. This corresponds to the immediate formation of a single condensate, which then hops over the set S as a random walk. In that case we can characterize the limit dynamics in a relatively simply way as the solution to a martingale problem for linear functions. In the more general case, in a first stage, several condensates will form, and interact via intermediate sites. Condensates do not split but can merge, and this coarsening dynamics eventually leads to a single condensate which then again moves over S as a random walk. In this case, we compute the generator explicitly making use of the separation of time scales and the martingales for the fast Wright-Fisher diffusion.

The rest of our paper is organized as follows. In Section 2 we introduce the model and present the main results and possible extensions. In Section 3 we collect general results on tightness and the comparison of semigroups and work out the simple two-site and fully connected cases. The proof of the general case is treated in Sections 4 and Sections 6 where we give a general definition of the limiting generator of a two-scale system in terms of projection operators on harmonic functions. In Section 5 we give

an explicit computation of the limiting generator in the case of 3 lattice sites, which already contains most of the difficulties of the general case.

2 Definitions and main results

2.1 The model and the auxiliary slow-fast system

We consider a finite set S , with associated jump rates $p(i, j) = p(j, i)$, $j, i \in S$ of a symmetric, irreducible random walk with $p(i, i) = 0$.

The symmetric inclusion process $(\eta(t) : t \geq 0)$ with parameter $m > 0$ based on $p(i, j)$ is then defined as in [2] to be the continuous-time Markov process on the configuration space \mathbb{Z}_+^S , with generator

$$\mathcal{L}_N^{SIP} f(\eta) = \sum_{i, j \in S} p(i, j) \eta_i \left(\frac{m}{2} + \eta_j \right) (f(\eta^{ij}) - f(\eta)) , \quad (2.1)$$

where $\eta_i \in \mathbb{Z}_+$ denotes the number of particles at site $i \in S$, and $\eta_k^{ij} = \eta_k - \delta_{i,k} + \delta_{j,k}$, $k \in S$ denotes the configuration where one particle moved from i to j . Since the total number of particles is conserved, we consider the process with generator (2.1) on the state space $\Omega_N = \{\eta \in \mathbb{Z}_+^S : \sum_{i \in S} \eta_i = N\}$ with a fixed total number $N \in \mathbb{N}$ of particles.

We consider $m = m_N$ dependent on the number of particles N such that $m_N \rightarrow 0$ as $N \rightarrow \infty$, which implies that diffusion becomes slower and the attractive inclusion interaction will create condensates. We are interested in the limiting dynamics of the rescaled process after accelerating time by a factor $\theta_N := \alpha/m_N$ with $\alpha > 0$. We will assume throughout the paper that

$$m_N \rightarrow 0, \quad Nm_N \rightarrow \infty \quad \text{so that} \quad \frac{\theta_N}{N} = \frac{\alpha}{m_N N} \rightarrow 0. \quad (2.2)$$

We denote by $e_i, i \in S$ the canonical unit vectors of $l_1(S)$ with entries $e_i(j) = \delta_{i,j}$. Consider the rescaled process $(x^N(t) : t \geq 0)$ defined by

$$x^N(t) := (\eta_i(\theta_N t)/N : i \in S).$$

This is a Markov process on the simplex $E = \{x \in [0, 1]^S : \sum_{i \in S} x_i = 1\}$ with generator

$$\mathcal{L}_N f(x) = \sum_{i, j \in S} \theta_N p(i, j) N x_i \left(\frac{m}{2} + N x_j \right) \left(f(x - \frac{1}{N} e_i + \frac{1}{N} e_j) - f(x) \right). \quad (2.3)$$

Assuming smooth f in (2.3), Taylor expansion of the right-hand side gives, using the symmetry of $p(i, j)$,

$$\begin{aligned} \mathcal{L}_N f(x) &= - \sum_{i, j \in S} \frac{\alpha}{4} p(i, j) (x_i - x_j) (\partial_{ij} f)(x) \\ &\quad + \frac{1}{2} \sum_{i, j \in S} p(i, j) x_i x_j \theta_N (\partial_{ij}^2 f)(x) + O(\theta_N/N), \end{aligned} \quad (2.4)$$

where we abbreviated

$$\partial_{ij} := \left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j} \right). \quad (2.5)$$

Here, by (2.2), the correction term denoted by $O(\theta_N/N)$ is a function of x which converges to zero uniformly in x , as $N \rightarrow \infty$. We will show that this correction terms can

be ignored and one can study the easier auxiliary process $(y^N(t) : t \geq 0)$ on E with generator

$$\begin{aligned} L_N &:= L + \theta_N L' \\ &= - \sum_{i,j \in S} \frac{\alpha}{4} p(i,j) (y_i - y_j) \partial_{ij} + \theta_N \frac{1}{2} \sum_{i,j \in S} p(i,j) y_i y_j \partial_{ij}^2. \end{aligned} \quad (2.6)$$

As $N \rightarrow \infty$ this is a slow-fast system with two scales, where the fast part

$$L' = \frac{1}{2} \sum_{i,j \in S} p(i,j) y_i y_j \partial_{ij}^2 \quad (2.7)$$

corresponds to a Wright-Fisher diffusion with absorbing states given by

$$\mathcal{A} := \{x \in E : p(i,j) x_i x_j = 0 \text{ for all } i, j \in S\}. \quad (2.8)$$

The slow part is a deterministic motion with generator

$$L = - \sum_{i,j \in S} \frac{\alpha}{4} p(i,j) (y_i - y_j) \partial_{ij}. \quad (2.9)$$

The limiting dynamics of the process with generator L_N is intuitively described as follows. The fast $(\theta_N L')$ -part of the generator pushes “infinitely” fast to the set of absorbing states \mathcal{A} , whereas the slow (L) part tries to move away from \mathcal{A} . To quantitatively describe this process, we first define the harmonic measure ν_x on \mathcal{A} for the Wright-Fisher diffusion $(X'(t) : t \geq 0)$ with generator L' . This is given by the hitting probabilities from any point $x \in E$

$$\nu_x(A) := \lim_{t \rightarrow \infty} \mathbb{P}_x(X'(t) \in A), \quad (2.10)$$

which are well defined for all measurable subsets $A \subset \mathcal{A}$. Of course we have $\nu_x = \delta_x$ for all $x \in \mathcal{A}$. In the spirit of [14], Theorem 1.11, we define the X' -harmonic projection operator

$$P : C(E, \mathbb{R}) \rightarrow \mathcal{H}(E) \quad \text{where} \quad Pf(x) := \int_{\mathcal{A}} f(a) \nu_x(da) \quad (2.11)$$

for all $f \in C(E, \mathbb{R})$, where the range of P is the set of L' -harmonic functions $\mathcal{H}(E)$. The function Pf is harmonic for L' and solves the Dirichlet problem $L'\psi = 0$ with $\psi(a) = f(a)$ for all $a \in \mathcal{A}$. In particular $Pf(a) = f(a)$ for $a \in \mathcal{A}$, and as a consequence $P^2 f = Pf$, i.e., P acts as a projection from the set $C(E, \mathbb{R})$ onto the set of harmonic functions. The limiting generator A of the two-scale system (2.6) is then given by $A = PL$, corresponding to infinitesimally moving away from \mathcal{A} under the slow dynamics with generator L and immediately being projected back onto \mathcal{A} . We will give our main results in the next subsection and a rigorous justification of this projection procedure in section 4.2.

2.2 Main results

Our main results rigorously characterize the limiting dynamics of the process $(x^N(t) : t \geq 0)$, by showing convergence to a limit process $(x(t) : t \geq 0)$ which concentrates on the subset $\mathcal{A} \subset E$. Furthermore, the set of corner points

$$\mathcal{C} := \{e_i : i \in S\} \subset \mathcal{A} \quad (2.12)$$

is in turn absorbing for the limit process, as we can conclude from the first theorem. In the following we speak of weak convergence on path space, if for every $T > 0$ the processes $(x^N(t), 0 \leq t \leq T)$ converge to $(x(t), 0 \leq t \leq T)$ weakly in the Skorokhod topology.

Theorem 2.1. 1. Assume that $x^N(0) \rightarrow x^0 \in \mathcal{C}$ in distribution. Then $(x^N(t) : t \geq 0)$ converges weakly on path space to $(x(t) : t \geq 0)$, which is a Markov chain on \mathcal{C} with initial condition $x(0) = x^0$ and generator

$$Af(e_i) = \sum_{j \in S} \frac{\alpha}{2} p(i, j) (f(e_j) - f(e_i)) . \quad (2.13)$$

2. In the fully connected case, i.e., if $p(i, j) > 0$ for all $i, j \in S$, then the same holds for a general initial condition. I.e., if $x^N(0) \rightarrow x^0 \in E$ in distribution, then $(x^N(t) : t \geq 0)$ converges weakly on path space to $(x(t) : t \geq 0)$, the process on \mathcal{C} with generator (2.13), and with initial condition $\mathbb{P}(x(0) = e_i) = \mathbb{E}[x_i^0]$.

So if the process is started from a configuration in \mathcal{C} , i.e., when (as $N \rightarrow \infty$) all the mass concentrates on a single site, all future configurations are of this type and the single pile of mass (condensate) performs a random walk on S with rates proportional to $p(i, j)$. The same holds in the fully connected case for general initial conditions, since in that case $\mathcal{A} = \mathcal{C}$ and only the corner points are absorbing for the Wright-Fisher part. For the initial condition in item 2, note that for non-random x^0 the harmonic measure is exactly $\nu_{x^0}(e_i) = x_i^0$, since the marginal dynamics on each site is given by a diffusion (see Section 3.3 for details) and only the corner points are absorbing.

For general symmetric $p(i, j)$, \mathcal{A} has a more complicated structure, the limit process is not ergodic on \mathcal{A} and there is no easy explicit formula for the harmonic measure ν_{x^0} . In the following we will describe the dynamics for general initial conditions under the additional assumption

$$p(i, j) \in \{0, 1\} \quad \text{for all } i, j \in S , \quad (2.14)$$

i.e., where edges are either connected or not with uniform jump rate 1. The more general case can be treated as well but is more complicated to formulate, and we will comment on that in a discussion in Section 2.3. Let

$$\hat{p}(i, j) = (1 - p(i, j)) \sum_{k \in S} p(i, k) p(k, j) \geq 0 \quad (2.15)$$

be the number of two-step connections between sites $i, j \in S$, which are not directly connected.

Theorem 2.2. Assume (2.14) and assume that at time zero $x^N(0) \rightarrow x^0 \in E$ in distribution. Then $(x^N(t) : t > 0)$ converges weakly on path space to $(x(t) : t \geq 0)$, which is a Markov process on \mathcal{A} with initial condition $x(0) \sim \mathbb{E}[\nu_{x^0}]$ determined by the harmonic measure (2.10) and generator

$$\begin{aligned} Af(x) = & \sum_{i, j \in S} \frac{\alpha}{2} \hat{p}(i, j) x_i x_j \partial_{ij}^2 f(x) \\ & + \sum_{j \in S} \delta_{x_j, 0} \left(\sum_{i \in S} \frac{\alpha}{2} p(i, j) x_i \right) \left(f \left(x + \sum_{i \in S} p(i, j) x_i (e_j - e_i) \right) - f(x) \right) . \end{aligned} \quad (2.16)$$

The generator of this process consists of two parts. The first part corresponds to an effective Wright-Fisher diffusion between sites which are connected by a path of length two only, which results from exchanging mass via the intermediate site. This mechanism of mass exchange is also studied heuristically in [8] for a totally asymmetric model. The second part corresponds to all the neighbouring mass of an empty site j accumulating on that site with a rate proportional to the total neighbouring mass. In the simplest case where there is only one neighbour with non-zero mass, this corresponds to a jump of a single pile. Therefore piles can exchange mass continuously or merge, but

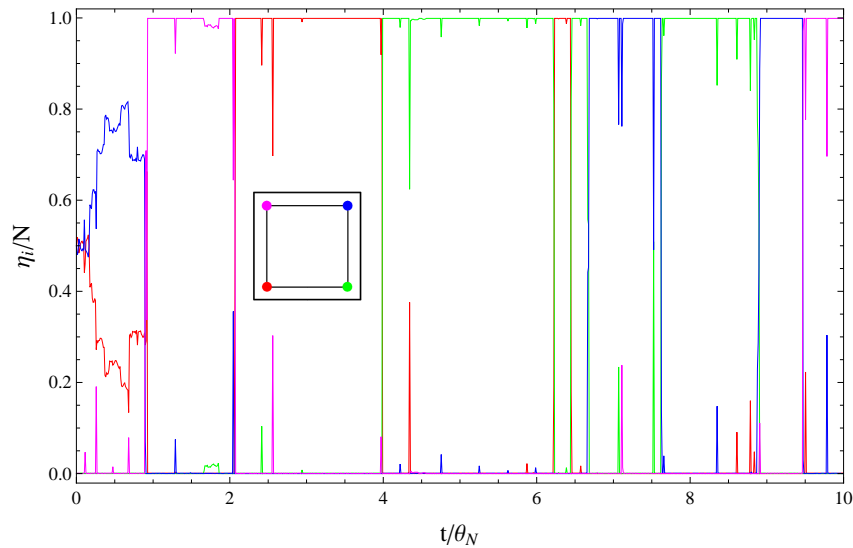


Figure 1: Simulation illustrating the main results by a sample path of the process with normalized occupation numbers η_i/N shown in different colors. The lattice is $S = \{1, \dots, 4\}$ with nearest neighbour connections and periodic boundary conditions. Initially, there is diffusive mass exchange between sites which are not directly connected, and later all mass concentrates on a single pile. Fluctuations are due to finite parameter values, where $N = 1000$, $m_N = 0.01$ and $\theta_N = 100$.

they never split and the set of corner points \mathcal{C} where all mass concentrates in a single pile is absorbing. For the inclusion process this absorbing set is reached on the same time scale θ_N as the stationary motion on the corner set takes place. This is different from other models such as the zero-range process, where these dynamics happen on different time scales [15, 16, 11, 12, 10].

Note that the initial condition for the limit process is not x^0 and convergence in the usual sense does not hold at time 0, since the process is “immediately” projected onto the absorbing set \mathcal{A} . The initial condition is then distributed as the absorption point on \mathcal{A} starting from x^0 , which is consistent with the right limit $x(0+)$.

2.3 Discussion

The general limit dynamics of Theorem 2.2 is illustrated in Figure 1 for a ring of 4 sites. After initial diffusive mass exchange between opposite, non-connected sites, the system reaches the absorbing set of corner points and turns into a jump process. Figure 2 illustrates the fully connected case (second item in Theorem 2.1) where the system instantaneously reaches a corner and a single pile performs a random walk on the lattice.

Our results hold for symmetric jump rates $p(i, j)$, since otherwise the Taylor expansion (2.4) would contain first order drift terms diverging with N , which would lead to infinite limiting speed of the clusters. Therefore the limit dynamics are not well defined unless the $p(i, j)$ are symmetric.

In the formulation of Theorem 2.2 we restrict ourselves to $p(i, j) \in \{0, 1\}$, only for simplicity of presentation. Our methods can be applied straightforwardly to the general case, only the concrete computations in Section 5 have to be adapted. Assuming $p(1, 2) = p(2, 1) = p$, $p(2, 3) = p(3, 2) = q$ and $p(1, 3) = p(3, 1) = 0$ for the case of 3 sites,

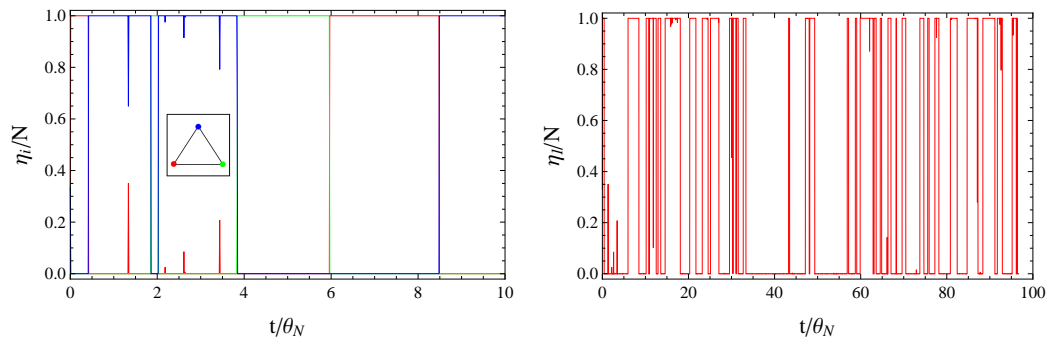


Figure 2: Simulation illustrating the main results for a fully connected lattice $S = \{1, 2, 3\}$ of three sites, where immediately all mass concentrates on a single pile. The left panel shows the evolution of the occupation number of all sites in different colors, and the right panel the evolution of site 1 for longer times. Fluctuations are smaller than in Fig. 1 since here $N = 10000$, $m_N = 0.001$ and $\theta_N = 1000$.

the continuous part of the limit generator A (2.16) then takes the form

$$\frac{\alpha}{4}(q-p)x_1x_3\partial_{13} + \frac{\alpha}{2}(qx_1+(p-q)x_1^2)(px_3+(q-p)x_3^2)\partial_{13}^2.$$

Compared to (5.14) and (2.16), there is an additional drift term and the diffusive term is no longer of Wright-Fisher type. This can be formulated for general lattices analogously. The jump part of A still takes the same form as in (2.16) on a general lattice.

3 A technical lemma and first results for the two-site and the fully connected case

This section consists of two parts.

1. In section 3.1, we show the following. If the auxiliary process $(y^N(t) : t \geq 0)$ with generator L_N (2.6) converges to a limiting process $(y(t) : t \geq 0)$ and if the “true process” $(x^N(t) : t \geq 0)$ with generator \mathcal{L}_N (2.3) converges along a subsequence, then this limit is the same process $(y(t) : t \geq 0)$. As a consequence, if the sequence $(x^N(t) : t \geq 0)$ is tight, then it has the limit characterized by the auxiliary process. This observation will be proved in Lemma 3.1 below and allows us to work with the simpler generator L_N to characterize the limit process.
2. In subsection 3.2 below we first consider the case of two sites, which is easier than the general case because the limiting process is a pure jump process on the set of unit vectors $\{e_1, e_2\}$. Then we work out the fully connected case, which is similar to the two site case, and where the limiting process is also a pure jump process on the set of unit vectors $\{e_i : i \in S\}$.

3.1 General results on comparison and tightness

We start with a technical lemma that we will apply later to justify dropping the $O(\theta_N/N)$ part in the Taylor expansion of the generator (2.4).

Lemma 3.1. *Suppose \mathcal{L}_N and L_N are generators of Markov processes on the compact state space E , such that on a common core of functions $\mathcal{K} \subset C(E)$ contained in continuous functions $C(E)$ we have*

$$\|\mathcal{L}_N f - L_N f\|_\infty \leq \|f\| \beta_N. \quad (3.1)$$

Here $\|\cdot\|_\infty$ denotes the supremum norm, β_N is a sequence of numbers converging to zero as $N \rightarrow \infty$, and $\|\cdot\|$ denotes a norm or seminorm such that for all $f \in \mathcal{K}$ and $t > 0$

$$\sup_N \sup_{0 \leq s \leq t} \|e^{sL_N} f\| < \infty. \quad (3.2)$$

Then we have:

1. For the associated semigroups, for all $t \geq 0$ and $f \in \mathcal{K}$

$$\lim_{N \rightarrow \infty} \|e^{tL_N} f - e^{t\mathcal{L}_N} f\|_\infty = 0. \quad (3.3)$$

2. If $\lim_{N \rightarrow \infty} e^{tL_N} f = S_t f$ for all $t \geq 0$, $f \in C(E)$ with $(S_t : t \geq 0)$ a Feller semigroup, then $\lim_{N \rightarrow \infty} e^{t\mathcal{L}_N} f = S_t f$ as well.
3. If the processes with generators \mathcal{L}_N and L_N are tight on the path space $D(0, \infty)$ of rcl functions (right-continuous with left limits), then they have the same limit process.

Proof. Pick $f \in \mathcal{K}$. Start from integration by parts (or Dyson's formula (see e.g. [17] p. 367) and use that $e^{s\mathcal{L}_N}$ is a contraction in the supremum norm to estimate

$$\begin{aligned} \|e^{tL_N} f - e^{t\mathcal{L}_N} f\|_\infty &= \left\| \int_0^t e^{s\mathcal{L}_N} (\mathcal{L}_N - L_N) e^{(t-s)L_N} f ds \right\|_\infty \\ &\leq \int_0^t \|(\mathcal{L}_N - L_N) e^{(t-s)L_N} f\|_\infty ds \\ &\leq \int_0^t \beta_N \|e^{(t-s)L_N} f\| ds \\ &\leq \beta_N t \sup_{0 \leq s \leq t} \|e^{sL_N} f\| \rightarrow 0 \end{aligned} \quad (3.4)$$

as $N \rightarrow \infty$ for all $t > 0$.

For the second item, let $f \in C(E)$, and for given $\epsilon > 0$ choose $g \in \mathcal{K}$ such that $\|g - f\|_\infty < \epsilon$. Then for N large enough we have, by item 1, $\|e^{tL_N} g - e^{t\mathcal{L}_N} g\|_\infty < \epsilon$ and by the contraction property of $e^{tL_N}, e^{t\mathcal{L}_N}$, $\|e^{tL_N}(f - g)\|_\infty, \|e^{t\mathcal{L}_N}(f - g)\|_\infty < \epsilon$. This implies $\|e^{tL_N} f - e^{t\mathcal{L}_N} f\|_\infty < 3\epsilon$ and if one of the terms converges to a semigroup, so does the other.

Finally, the third item follows because a Markov process on the path space $D(0, \infty)$ is uniquely determined by its semigroup. \square

Lemma 3.2. *The sequence of Markov processes $(x^N(t) : t > 0)$ on a complete separable metric space E is tight on the path space $D(0, \infty)$ of rcl functions if*

1. The sequence $(x^N(t) : t \geq 0)$ is stochastically bounded on $D[0, \infty)$, and
2. For all $f \in C_c^\infty(E)$, denoting compact smooth functions (which are uniformly dense in $C(E)$) and all $\epsilon > 0$ there exist non-negative random variables $Z_N(\delta, f, \epsilon)$ such that for all δ

$$\left| \mathbb{E}[f(x^N(t+u)) - f(x^N(t)) | x^N(t) = x] \right| \leq \mathbb{E}[Z_N(\delta, f, \epsilon) | x^N(t) = x] \quad (3.5)$$

with probability 1 for all $t \geq \epsilon$ and $0 \leq u \leq \delta$, and

$$\lim_{\delta \searrow 0} \limsup_{N \rightarrow \infty} \mathbb{E}[Z_N(\delta, f, \epsilon)] = 0. \quad (3.6)$$

The conditions are sufficient but not necessary, for a proof, see e.g. [19], Lemma 3.11 or [20], Chapter 3. The first condition will always be fulfilled since our state space E is a compact, finite dimensional simplex and $\|x^N(t)\|_\infty \leq 1$. The second condition has been modified for our purposes, and ensures boundedness of the modulus of continuity of sample paths starting the process from any strictly positive time and will be verified later. At time 0 the sequence of processes we consider is actually not tight, since in the limit it gets instantly projected onto an absorbing state of the Wright-Fisher diffusion L' . So we can show convergence only for all positive times $t > 0$, and by right-continuity of sample paths we can extend the limit process uniquely to all $t \geq 0$. We will show in Section 4 that this extension is consistent with the harmonic measure, describing the absorption of the fast Wright-Fisher diffusion.

3.2 The two site case

We start now with the simplest case

$$S = \{1, 2\} \quad \text{and} \quad p(1, 2) = p(2, 1) = 1, \quad (3.7)$$

where S contains only two sites and $\mathcal{A} = \mathcal{C} = \{e_1, e_2\}$.

Theorem 3.3. *Assume (3.7) with initial conditions $x^N(0)$ such that $x^N(0) \rightarrow x^0 \in [0, 1]$ in distribution as $N \rightarrow \infty$. Then the process $(x^N(t) : t > 0)$ converges weakly on path space to $(x(t) : t \geq 0)$, which is a simple random walk on $\{e_1, e_2\}$ with jump rates $\alpha/2$ and initial condition $x(0) \in \{e_1, e_2\}$ with $\mathbb{P}(x(0) = e_1) = x_1^0$.*

The rest of this section is devoted to the proof. Since S consists of two elements we abbreviate $x_1 = x$, and because in the process $x_1 + x_2 = 1$ is conserved, we effectively have only one variable (denoted by x). The Taylor expansion leading to (2.4) then gives in terms of x :

$$\begin{aligned} \mathcal{L}_N f(x) &= \left(-\frac{1}{2} \theta_N m_N (2x - 1) \partial_x + 2 \theta_N x(1 - x) \partial_x^2 \right) f(x) + O(\theta_N/N) \\ &=: L_N f(x) + O(\theta_N/N) \end{aligned} \quad (3.8)$$

where the terms $O(\theta_N/N)$ go to zero uniformly in x , as $N \rightarrow \infty$. The generator

$$L_N = -\frac{1}{2} \alpha (2x - 1) \partial_x + 2 \theta_N x(1 - x) \partial_x^2 \quad (3.9)$$

appearing in (3.8) is a Wright-Fisher diffusion with mutation where in front of the diffusion term $x(1 - x) \partial_x^2$ we have a factor θ_N , i.e., the diffusion part (also called “genetic drift part” in the population dynamics language) of the process is accelerated w.r.t. the mutation process.

This implies intuitively that the process will tend to evolve “immediately” towards the fixed points of the genetic drift process, which are the homo-zygotes $x = 1$ and $x = 0$, and the mutation process will then lead to a random flipping between these two states.

First, we show that we can apply Lemma 3.1 to our case. Since this is slightly technical, we state it as a separate lemma.

Lemma 3.4. *The generators \mathcal{L}_N and L_N of (3.8) satisfy the assumptions of Lemma 3.1 with the seminorm $\|f\| = (\|f''\|_\infty + \|f'''\|_\infty)$ and with \mathcal{K} the set of polynomials as a common core.*

Proof. The approximate generator reads (for simplicity we put $\alpha/2 = 1$ in the following)

$$L_N = -(2x - 1) \partial_x + 2 \theta_N x(1 - x) \partial_x^2$$

and for the difference with the true generator we have the estimate

$$\|\mathcal{L}_N f - L_N f\|_\infty \leq C \left(\frac{1}{N} \|f''\|_\infty + \frac{\theta_N}{N} \|f'''\|_\infty \right) \leq \frac{C\theta_N}{N} (\|f''\|_\infty + \|f'''\|_\infty)$$

from a standard expression of the remainder in Taylor expansions. The seminorm $\|f\| = (\|f''\|_\infty + \|f'''\|_\infty)$ can thus be used in (3.1).

So in order to see that (3.2) holds for this choice of norm, it is sufficient to prove that for the auxiliary process $y^N(t)$ with generator L_N , the expectations

$$\mathbb{E}_y f(y^N(s)) = e^{sL_N} f(y)$$

have bounded second and third derivatives in x , uniformly in $0 \leq s \leq t$ and $N \in \mathbb{N}$, for sufficiently many f . Choosing for f the polynomial $f_n(y) = y^n$, we have,

$$L_N f_n = 2\theta_N n(n-1)(f_{n-1} - f_n) + n(f_{n-1} - f_n) - n f_n$$

This can be interpreted as follows: the first two terms form the generator of a Markov chain $(n_t : t \geq 0)$ that jumps from n to $n-1$ at rate $2\theta_N n(n-1) + n$ corresponding to the classical ancestral dual of Wright-Fisher diffusion (Kingman's coalescent + mutation contribution), and the last term is an extra killing part. For more details on that connection see e.g. the nice account on duality in population models [18]. Using the Feynman-Kac formula we then obtain,

$$\psi(N, t, n, y) := \mathbb{E}_y f_n(y^N(t)) = \mathbb{E}_n^N (e^{-\int_0^t n_s ds} f_{n_t}(y))$$

where \mathbb{E}_n^N denotes expectation for the process on \mathbb{N} with generator

$$K_N g(n) = \theta_N n(n-1)(g(n-1) - g(n)) + n(g(n-1) - g(n)) .$$

As a consequence, since the process with generator K_N only lowers the starting n in the course of time, we obtain the uniform estimate

$$\left| \frac{d^k}{dx^k} \psi(N, t, n, y) \right| = \left| \mathbb{E}_n^N (e^{-\int_0^t n_s ds} y^{n_t-k} n_t(n_t-1) \dots (n_t-k+1)) \right| \leq n^k$$

for all $y \in [0, 1]$, which gives (3.2). Finally, it is clear that on the compact simplex E the polynomials form a common core of the generators L_N, \mathcal{L}_N . \square

So Lemma 3.1 guarantees that we can identify the limit from the auxiliary process $(y^N(t) : t \geq 0)$ with generator L_N . We first show that the limit process concentrates on the set $\{0, 1\}$ and then characterize it as a solution to the martingale problem of the claimed limiting process.

Lemma 3.5. *We have for all $t > 0$*

$$\sup_{y \in [0,1]} \mathbb{E}_y [y^N(t)(1 - y^N(t))] \rightarrow 0 \quad \text{as } N \rightarrow \infty, \quad (3.10)$$

and furthermore

$$\limsup_{N \rightarrow \infty} \theta_N \sup_{y \in [0,1]} \mathbb{E}_y [y^N(t)(1 - y^N(t))] \leq C \quad (3.11)$$

for some $C > 0$. The same holds for $x^N(t)(1 - x^N(t))$ of the original process with generator \mathcal{L}_N .

Proof. We abbreviate $f(y) = 2y(1 - y)$, $g(y) = 2(2y - 1)^2$ and compute

$$\frac{d}{dt} \mathbb{E}_y f(y^N(t)) = \mathbb{E}_y(g(y^N(t))) - 2\theta_N \mathbb{E}_y f(y^N(t)) .$$

Putting $\psi_t(y) = \mathbb{E}_y f(y^N(t))$, $\varphi_t(y) = \mathbb{E}_x g(y^N(t))$, we rewrite this equation as

$$\frac{d}{dt} \psi_t = \varphi_t - 2\theta_N \psi_t ,$$

which has the standard form of a slow-fast system (see e.g. [13] for more background), omitting the fixed argument y . This can be rewritten as

$$\psi_t = \psi_0 e^{-2\theta_N t} + \frac{e^{-2\theta_N t}}{\theta_N} \int_0^{\theta_N t} \varphi_{s/\theta_N} e^{2s} ds . \quad (3.12)$$

Using now the fact that $y \in [0, 1]$, we obtain that $\varphi \leq 2$ and $\psi \leq 1$ which gives

$$0 \leq \psi_t(y) \leq e^{-\theta_N t} + \frac{2}{\theta_N} (1 - e^{-2\theta_N t}) \quad (3.13)$$

which goes to zero as $N \rightarrow \infty$ for all $t > 0$, thus proving (3.10). The estimate (3.13) also implies the second statement (3.11).

The same argument holds for $x^N(t)(1 - x^N(t))$ if we replace φ_t by $\varphi_t + O(\theta_N/N)$ including the remainder terms, since we only use boundedness of φ_t for N large enough. \square

This shows that the limit process (if it exists) concentrates on the state space $\{0, 1\}$ for all $t > 0$, and the estimate can also be used to show tightness using the criterion of Lemma 3.2, which implies existence of a subsequential limit by Prohorov's theorem (see e.g. [20], Chapter 3).

Lemma 3.6. *The processes $(y^N(t) : t > 0)$ and $(x^N(t) : t > 0)$ are tight on the path space $D(0, \infty)$.*

Proof. As mentioned before, the state space $E = [0, 1]$ is compact and the first condition of Lemma 3.2 holds clearly. For the second condition, we first note that the process is not tight for $t = 0$ since it is immediately projected onto the absorbing set \mathcal{A} which is $\{0, 1\}$ in this simple case. For all $t > 0$ we can show tightness for the auxiliary process first by noting that for any smooth function f on $[0, 1]$

$$\mathbb{E}[f(y^N(t+u)) - f(y^N(t)) | y^N(t) = y] = \mathbb{E}\left[\int_t^{t+u} L_N f(y^N(s)) ds \Big| y^N(t) = y\right] . \quad (3.14)$$

On the right-hand side

$$L_N f(y) = -(2y - 1)f'(y) + \theta_N y(1 - y)f''(y) \quad (3.15)$$

and the first term is bounded by $u \|f'\|_\infty$ since $y \in [0, 1]$. For the second term we can exchange expectation and time integration since the integrand is positive and bounded above, and use (3.12) to get

$$\begin{aligned} & \theta_N \mathbb{E}\left[\int_t^{t+u} y^N(s)(1 - y^N(s)) ds \Big| y^N(t) = y\right] \\ &= \int_t^{t+u} \left(\theta_N y(1 - y) e^{-2\theta_N(s-t)} + e^{-2\theta_N(s-t)} \int_0^{\theta_N(s-t)} \varphi_{w/\theta_N} e^{2w} dw \right) ds \\ &\leq \frac{1}{2} y(1 - y) (1 - e^{-2\theta_N u}) + 2u , \end{aligned} \quad (3.16)$$

where we also used boundedness of φ as before. This is independent of t and for all $t \geq \epsilon$ and $0 \leq u \leq \delta$ we can therefore use

$$Z_N(\delta, f, \epsilon) := \delta \|f'\|_\infty + \left(\frac{1}{2}y(1-y) + 2\delta\right) \|f''\|_\infty \quad (3.17)$$

to bound the modulus of (3.14) as in Lemma 3.2. For every fixed $\epsilon > 0$ this bound holds uniformly in N , and uniformly in the conditional value $y \in E$ at time t . Using that $y = y^N(t)$ for some $t \geq \epsilon$ it follows from Lemma 3.5 that for all $y_0 \in E$

$$\mathbb{E}_{y_0}[Z_N(\delta, f, \epsilon)] \rightarrow 0 \quad (3.18)$$

as $N \rightarrow \infty$ and $\delta \searrow 0$, since f is smooth and all derivatives are bounded on $[0, 1]$. This implies tightness by Lemma 3.2. Since $\epsilon > 0$ was arbitrary we have tightness for all strictly positive times $t > 0$.

The same argument holds for the original process $(x^N(t) : t > 0)$, since also third derivatives of f are bounded and additional terms are of the form $\delta \|f'''\|_\infty (C + \theta_N/N)$. \square

In order to identify the limit process by a martingale problem, the following well-known result gives a simplified version for a process that jumps with rate λ between 0 and 1.

Lemma 3.7. *Let X_t be a Markov process with values in $\{0, 1\}$ and $\lambda > 0$. If*

$$M_t := X_t - X_0 - \int_0^t \lambda(2X_s - 1) ds \quad (3.19)$$

is a martingale, then X_t is the process that jumps with rate λ between 1 and 0.

Proof. Denote by

$$Lf(x) = \lambda x(f(0) - f(1)) + \lambda(1 - x)(f(1) - f(0)) = -\lambda(2x - 1)(f(1) - f(0))$$

the generator of the process jumping back and forth between 0, 1 at rate λ . Let $f : \{0, 1\} \rightarrow \mathbb{R}$, then $f(x) = xf(1) + (1 - x)f(0) = x(f(1) - f(0)) + f(0)$ is always linear. Therefore, using (3.19)

$$f(X_t) - f(X_0) - \int_0^t Lf(X_s) ds = M_t^f \quad (3.20)$$

is a martingale for all $f : \{0, 1\} \rightarrow \mathbb{R}$, which characterizes the process. \square

Now we can conclude the proof of Theorem 3.3. Applying the generator L_N (3.8) of the auxiliary process to the function $f(y) = y$ gives that

$$M^N(t) := y^N(t) - y^N(0) - \frac{\alpha}{2} \int_0^t (2y^N(s) - 1) ds$$

is a martingale. Therefore, if $(y^N(t) : t > 0) \rightarrow (x(t) : t > 0)$ along a subsequence, for the limiting process we have that

$$M(t) := x(t) - x(0) - \frac{\alpha}{2} \int_0^t (2x(s) - 1) ds$$

is a martingale. Thus the limit is the claimed jump process on $\{0, 1\}$. By Lemma 3.1 this is also the limit of the original process $x^N(t)$.

Since the limit process $(x(t) : t > 0)$ has right-continuous paths the consistent initial condition is given by

$$x(0) := \lim_{\epsilon \searrow 0} \lim_{N \rightarrow \infty} y^N(\epsilon), \quad (3.21)$$

which concentrates on $\{0, 1\}$ by Lemma 3.5. To see that this has the distribution claimed in Theorem 3.3, denote $\varphi(t, N) = \mathbb{E}_{y^N(0)}(y^N(t))$. Then using the generator (3.9), we see that

$$\frac{d\varphi(t, N)}{dt} = -\frac{\alpha}{2}(2\varphi(t, N) - 1)$$

which gives

$$\varphi(t, N) = y^N(0) + \frac{1}{2}(1 - e^{-\alpha t}).$$

Taking the limits $N \rightarrow \infty$ and then $t \downarrow 0$, and using that the limiting process is concentrated on $\{0, 1\}$ gives the initial condition as claimed.

Remark 3.8. We will see in Section 4 that in general the limiting semigroup $S(t)$ has the property that $S(0) = P$ is a projection onto harmonic functions on E . So if μ is the distribution of limiting initial condition x^0 as given in Theorems 2.1 and 2.2, the initial distribution of $y(0)$ for the limit process is given by $\mu P = \mathbb{E}_\mu[\nu_{x^0}]$ since $(Pf)(x) = \mathbb{E}_{\nu_x}[f]$ for all $f \in C(E, \mathbb{R})$. For a symmetric diffusion on $[0, 1]$ the harmonic measure ν_x on $\{0, 1\}$ is then given by $\nu_x = x_1^0$, i.e. equal to the first component of $x^0 \in E$ (see the second item of Theorem 2.1).

3.3 The fully connected case

We will see in the next section that Lemma 3.1, the concentration Lemma 3.5 and the tightness argument in Lemma 3.6 can be extended straightforwardly to the general case. In case the set S is fully connected, i.e. all $p(i, j) > 0$, we have $\mathcal{A} = \mathcal{C}$ and the limit process still concentrates on the corner points and can be characterized by a simple martingale problem. This will provide a proof of the second statement of Theorem 2.1.

The following lemma gives the martingale characterization of the process jumping with rates $p(i, j)$ from e_i to e_j .

Lemma 3.9. Let $(y(t) : t \geq 0)$ be a process on $\{e_i, i \in S\}$, and denote $y_i(t) = y(t) \cdot e_i = \mathbb{1}(y(t) = e_i)$. If for all i

$$M^i(t) := y_i(t) - y_i(0) - \sum_{j \in S} \int_0^t (y_j(s) - y_i(s)) p(i, j) ds \quad (3.22)$$

is a martingale, then y is the random walk jumping at rate $p(i, j) = p(j, i)$ between e_i and e_j .

Proof. As this is a straightforward extension of Lemma 3.7, we will be brief. On the corner points \mathcal{C} of the finite simplex E a general function

$$f(y) = \sum_{i \in S} f(e_i) y_i \quad (3.23)$$

can again be written as a linear combination of simple linear functions y_i , which implies the simpler martingale characterization along the same arguments as Lemma 3.7 in the two-site case. \square

The generator $L_N = \sum_{i, j \in S} L_N^{ij}$ (2.6) applied to the linear function $f_k(y) = y_k$ gives

$$L_N f_k(y) = \sum_{i \in S} \frac{\alpha}{2} p(k, i) (x_i - x_k)$$

Therefore, for all $k \in S$

$$y_k^N(t) - y_k^N(0) - \sum_{i \in S} \int_0^t \frac{\alpha}{2} p(k, i) (y_i^N - y_k^N)$$

is a martingale. So the limit along any convergent subsequence of $(y^N(t) : t > 0)$ will fulfill the condition in Lemma 3.9 and is therefore unique. Again, the initial conditions of the right-continuous limit process can be defined as in (3.21) and have distribution $\mathbb{E}[\nu_{x^0}]$ by the same arguments, which proves the second statement of Theorem 2.1.

4 The general case

4.1 Concentration and tightness

Recall the generators (2.4) and (2.6) of the original and the auxiliary process,

$$\begin{aligned} \mathcal{L}_N f(x) &= - \sum_{i,j \in S} \frac{\alpha}{4} p(i, j) (x_i - x_j) (\partial_{ij} f)(x) \\ &\quad + \frac{1}{2} \sum_{i,j \in S} p(i, j) x_i x_j \theta_N (\partial_{ij}^2 f)(x) + O(\theta_N/N) \\ &= Lf(x) + \theta_N L'f(x) + O(\theta_N/N). \end{aligned} \quad (4.1)$$

Individual terms for fixed i, j are precisely of the form dealt with in the two-site case. Therefore, Lemmas 3.5 and 3.6 on concentration and tightness are relatively straightforward to generalize, and we only give short proofs in the following.

Lemma 4.1. *We have for all $t > 0$ for the auxiliary process*

$$p(i, j) \sup_{y \in E} \mathbb{E}_y [y_i^N(t) y_j^N(t)] \rightarrow 0 \quad \text{as } N \rightarrow \infty, \quad (4.2)$$

and furthermore

$$p(i, j) \limsup_{N \rightarrow \infty} \theta_N \sup_{y \in [0,1]} \mathbb{E}_y [y_i^N(t) y_j^N(t)] \leq C \quad (4.3)$$

for some $C > 0$. The same holds for the original process $(x^N(t) : t > 0)$ with generator \mathcal{L}_N .

Proof. Fix $i, j \in S$ with $p(i, j) = p(j, i) > 0$ (otherwise there is nothing to prove). Analogous to the proof of Lemma 3.5 we can compute

$$L'(y_i y_j) = -2p(i, j) y_i y_j \quad (4.4)$$

and

$$L(y_i y_j) = -\frac{\alpha}{4} \left(y_i \sum_{k \in S} p(j, k) (y_j - y_k) + y_j \sum_{k \in S} p(i, k) (y_i - y_k) \right). \quad (4.5)$$

The second term is bounded in absolute value by the constant

$$C := \alpha \max_{i,j \in S} \sum_{k \in S} (p(i, k) + p(j, k))$$

and only the first term is relevant, since it is multiplied by θ_N . Using the notation

$$\varphi_t(y) = \mathbb{E}_y [y_i^N(t) y_j^N(t)] \quad \text{and} \quad \psi_t(y) = \mathbb{E}_y [L(y_i^N y_j^N)] \quad (4.6)$$

we get the estimate

$$0 \leq \psi_t(y) \leq e^{-2p(i,j)\theta_N t} + \frac{C}{2p(i,j)\theta_N} (1 - e^{-2p(i,j)\theta_N t}) \quad (4.7)$$

with the same arguments that lead to (3.13). The rest of the proof follows analogously to the two-site case. \square

Lemma 4.2. *The sequences $(x^N(t) : t > 0)$ for the full process and $(y^N(t) : t > 0)$ for the auxiliary process are tight on the path space $D(0, \infty)$ of rcll functions.*

Proof. We use again the criterion in Lemma 3.2 on the compact state space E and the same argument as in the two-site case. The only new element is the estimate of the generator which follows analogously using

$$\sup_{y \in E} \mathbb{E}_y [L_N f(y^N(t))] \leq \sum_{i,j \in S} \left(p(i,j) \|\partial_{ij} f\|_\infty + C \|\partial_{ij}^2 f\|_\infty \right) \quad (4.8)$$

since all $y_i^N(t) \in [0, 1]$ and for $t > 0$ we can use the estimate Lemma 4.1. Again, f is smooth on the compact space E so that all partial derivatives are bounded. \square

So the limiting process exists for all positive times $t > 0$ and concentrates on the absorbing set \mathcal{A} of the fast Wright-Fisher diffusion. The initial condition on \mathcal{A} will be given consistently by right limits of the sample paths which coincide with the absorption probabilities of the Wright-Fisher diffusion, as we will see in the next subsections.

4.2 Slow-fast Markovian systems and projection of Markov processes

In the general case, the generator (2.6) still consists of a fast Wright-Fisher diffusion part and a slow mutation drift part. The limiting process will take place on the absorbing set \mathcal{A} of the Wright-Fisher diffusion, which in the case of general $p(i, j)$ is a richer set than the set of corner points $\mathcal{C} = \{e_i : i \in S\}$.

We first describe the limit motion of a general two-scale Markovian system in greater generality, following closely the results in [14]. Let us consider $(X(t) : t \geq 0)$, $(X'(t) : t \geq 0)$ Markov processes on a compact metric space E with generators L, L' . Both generators should have a common core \mathcal{K} (which is given by the smooth functions $C^\infty(E)$ in our case) and the corresponding semigroups e^{tL} and $e^{tL'}$ are defined on the Banach space $C(E, \mathbb{R})$ of continuous functions with the supremum norm $\|\cdot\|_\infty$. Suppose X'_t has a Borel measurable set of absorbing states $\mathcal{A} \subset E$ such that the hitting probabilities from any point $x \in E$ are well defined and given by the harmonic measure ν_x as defined in (2.10). Recall also the definition (2.11) of the projection operator P on the set of L' -harmonic functions $\mathcal{H}(E)$. Denote by $S_\kappa(t)$ the semigroup of the two-scale process X^κ with generator $L + \kappa L'$. The following result describes the limiting process as $\kappa \rightarrow \infty$ by a semigroup on the harmonic functions.

Theorem 4.3. *Suppose that $e^{tL'} f \rightarrow P f$ as $t \rightarrow \infty$ for all $f \in C(E, \mathbb{R})$ and that the operator*

$$A : C(E, \mathbb{R}) \rightarrow \mathcal{H}(E) \quad \text{where} \quad A f := P L f \quad (4.9)$$

generates a Markov semigroup $S(t)$ on $\mathcal{H}(E)$ (the range of P). Then $S_\kappa(t) f \rightarrow S(t) f$ as $\kappa \rightarrow \infty$ for all $f \in \mathcal{H}(E)$ and $t \geq 0$. The limiting process generated by A concentrates on the absorbing set \mathcal{A} and has initial condition given by the harmonic measure ν_{x^} if $X^\kappa(0) \rightarrow x^*$ for some $x^* \in E$.*

Proof. Convergence of the semigroups to the semigroup with generator A follows from [14], Theorem 2.1. The only condition to check is $\text{dom}(A) \subset \overline{\mathcal{R}(\lambda - A)}$ for some $\lambda > 0$. This condition is clearly satisfied if A generates a Markov process, because in that case $\mathcal{R}(\lambda - A)$ equals $C(\mathcal{A}, \mathbb{R})$.

With Dyson's formula ([17] p. 367) we have for all f in the core \mathcal{K}

$$S_\kappa(t) f - e^{\kappa t L'} f = \int_0^t e^{\kappa(t-s)L'} L S_\kappa(s) f ds. \quad (4.10)$$

Using convergence of the semigroups $S_\kappa(t)$ and $e^{\kappa t L'}$ as $\kappa \rightarrow \infty$ we get

$$S(t)f - Pf = \int_0^t A S(s)f ds. \quad (4.11)$$

This shows that A generates $S(t)$ and in particular that $S(0) = P$. For harmonic $f \in \mathcal{H}(E)$ we have $S(0)f = Pf = f$ and $S(t)$ is indeed a semigroup on $\mathcal{H}(E)$. This implies that the limit process generated by A concentrates on the absorbing set \mathcal{A} as explained in more detail below, and that by the definition of the projection P (2.11) the initial condition of the process is given by the harmonic measure ν_x . \square

Every L' -harmonic function on E is uniquely determined by its boundary values on \mathcal{A} (cf. (2.11)). So in addition to the projection P we can define the *harmonic extension*

$$\mathcal{P} : C(\mathcal{A}, \mathbb{R}) \rightarrow \mathcal{H}(E) \quad \text{where} \quad \mathcal{P}f(x) := \int \nu_x(da) f(a). \quad (4.12)$$

Notice that P and \mathcal{P} are defined by the same formula, but act on different spaces. In particular we have $Pf = (\mathcal{P}(f|_{\mathcal{A}}))$, i.e., the harmonic projection is the same as the harmonic extension of the restriction of f to \mathcal{A} . With this notation we can define

$$A' : C(\mathcal{A}, \mathbb{R}) \rightarrow C(\mathcal{A}, \mathbb{R}) \quad \text{where} \quad A'f := (A\mathcal{P}f)|_{\mathcal{A}}, \quad (4.13)$$

which highlights the fact that the limit semigroup generated by A actually defines a process on the absorbing set \mathcal{A} . Via the identification of $\mathcal{H}(E)$ with $C(\mathcal{A}, \mathbb{R})$ the operators A and A' both describe infinitesimal motion away from \mathcal{A} due to L , instantaneously followed by projection onto \mathcal{A} . Formally, in order to define A' on $f \in C(\mathcal{A}, \mathbb{R})$ and act with L we first have to extend the function f to the whole configuration space E : therefore we first harmonically extend f (action of \mathcal{P}), then apply L , then project (action of P) and restrict again to \mathcal{A} .

In order to apply Theorem 4.3 we have to assure that the operator A' generates a Markov process. This will be done in the next Section and below, where we explicitly compute A' and recognize it as the generator of a process on \mathcal{A} with a jump and a diffusion part. The condition of convergence of the semigroup of the fast diffusive part of the generator in our case is covered by the following result.

Proposition 4.4. *Let $L' = \frac{1}{2} \sum_{i,j \in S} p(i,j) y_i y_j \partial_{ij}^2$ be the generator (2.7) of the Wright-Fisher diffusion X' with absorbing states \mathcal{A} (2.8). Then*

$$\|e^{tL'}f - Pf\|_\infty \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (4.14)$$

Proof. Using that P is the projection on the absorbing set we have

$$(e^{tL'} - P)f(x) = \mathbb{E}_x[f(X'(t)) - f(X'(\infty))].$$

Denoting by τ the time of absorption of the Wright-Fisher diffusion, we get

$$\begin{aligned} \mathbb{E}_x[f(X'(t)) - f(X'(\infty))] &= \mathbb{E}_x[(f(X'(t)) - f(X'(\tau))) \mathbb{1}(\tau > t)] \\ &\leq 2\|f\|_\infty \mathbb{P}_x[\tau > t]. \end{aligned} \quad (4.15)$$

This converges to zero since f is bounded on the compact state space E , and the Wright-Fisher diffusion gets absorbed almost surely in finite time with a uniformly bounded mean absorption time (see e.g. [21]). \square

To illustrate how the projected generator A' of (4.13) looks like in a concrete case, let us look back at the example of two sites with $E = [0, 1]$, studied in the previous section. The X' -process is the Wright-Fisher diffusion with generator $L' = 2x(1-x)\partial_x^2$, with absorbing states 0 and 1. The X process is the deterministic process with generator $L = -\frac{\alpha}{2}(2x-1)\partial_x$. Since $X'(t)$ is a martingale, stopping at the time of absorption gives that the harmonic measure equals

$$\nu_x(\{1\}) = 1 - \nu_x(\{0\}) = x.$$

As a consequence, for a function $f : \{0, 1\} \rightarrow \mathbb{R}$, its harmonic extension is simply given by linear interpolation: $\mathcal{P}f(x) = xf(1) + (1-x)f(0)$ and therefore

$$(L\mathcal{P}f)(x) = -\frac{\alpha(2x-1)}{2} \frac{d}{dx} (xf(1) + (1-x)f(0)) = -\frac{\alpha(2x-1)}{2} (f(1) - f(0)).$$

Thus

$$A'f(x) = (PL\mathcal{P})f(x) = \frac{\alpha}{2} [x(f(0) - f(1)) + (1-x)(f(1) - f(0))]$$

which in the case of $\frac{\alpha}{2} = 1$ takes values

$$A'f(1) = f(0) - f(1) \quad \text{and} \quad A'f(0) = f(1) - f(0).$$

So A is the generator of the process which flips at rate 1 between the states of the absorbing set $\mathcal{A} = \{0, 1\}$, as we found in Section 3.2 using a martingale characterization.

5 Computation of the generator A : the case of three sites

In the case of general connections $p(i, j)$ and general finite sets S , the possible limiting behaviors of the process is more complicated. Therefore we focus on the case of Theorem 2.2 with $p(i, j) \in \{0, 1\}$ and compute the generator A (4.13) using that

$$PL = P \lim_{h \searrow 0} \left(\frac{e^{hL} - I}{h} \right) = \lim_{h \searrow 0} P \left(\frac{e^{hL} - I}{h} \right) \quad (5.1)$$

by continuity of the projection P . Note that by Theorem 4.3 it suffices to compute A rather A' to characterize the limit process. We start here with a system of three sites with closed boundaries and nearest neighbor connections. This case already contains most of the interesting aspects of the limiting dynamics which now has both a diffusive and a pure jump part.

5.1 Generator and absorbing set

We consider the three site case with $p(1, 2) = p(2, 1) = p(2, 3) = p(3, 2) = 1$ and $p(1, 3) = p(3, 1) = 0$. The generator (2.6) is

$$\begin{aligned} L_N &= -a(x_1 - x_2)\partial_{12} - a(x_2 - x_3)\partial_{23} \\ &+ \frac{\theta_N}{2}x_1x_2\partial_{12}^2 + \frac{\theta_N}{2}x_2x_3\partial_{23}^2 \end{aligned} \quad (5.2)$$

where we write $a = \alpha/4$ in the following to simplify notation. Recall the notation from (2.6) for this case: the “fast Wright-Fisher diffusion” has generator

$$L' = \frac{1}{2}x_1x_2\partial_{12}^2 + \frac{1}{2}x_2x_3\partial_{23}^2, \quad (5.3)$$

whereas, the “slow deterministic” part has generator

$$L = -a(x_1 - x_2)\partial_{12} - a(x_2 - x_3)\partial_{23}. \quad (5.4)$$

The process $x'(t)$ with generator L' is a Wright-Fisher diffusion, for which, $x'_1(t)$, $x'_2(t)$, $x'_3(t)$ and $x'_1(t)x'_3(t)$ are martingales since $L'x_1 = L'x_2 = L'x_3 = L'(x_1x_3) = 0$. Its absorbing set is

$$\mathcal{A} = (0, z, 0) \cup \{(u, 0, w) : u + w = z\} \quad (5.5)$$

given by a point and a line segment. Here we assume that the total mass in the system is $z \leq 1$, for our results to generalize more easily to larger systems.

5.2 Dynamics on the line segment: diffusion and jumps

We start with a point on the line segment $x(0) = (u, 0, w)$ and need to compute (see (5.1)):

$$Af(x(0)) = \lim_{h \searrow 0} \frac{1}{h} \left(\mathbb{E}_{x(h)}[f(x'(\tau))] - f(u, 0, w) \right) \quad (5.6)$$

where $x(h) = (u(1 - ah), zah, w(1 - ah)) \in E \setminus \mathcal{A}$ is the initial condition for the Wright-Fisher diffusion $x'(t)$. This initial condition comes from starting from the point $(u, 0, w)$ from the absorbing set and evolving according to the deterministic part of the dynamics a small amount of time h .

Then, since $x'_2(\tau) \in \{0, z\}$ takes only two values, we have

$$\begin{aligned} \mathbb{E}_{x(h)}[f(x'(\tau))] &= f(0, z, 0) \mathbb{P}_{x(h)}[x'_2(\tau) = z] \\ &\quad + \mathbb{E}_{x(h)}[f(x'(\tau)) \mathbb{1}(x'_2(\tau) = 0)] . \end{aligned} \quad (5.7)$$

Since $x'_2(t)$ is a martingale on $[0, z]$, the probability in the first line is given by ah and we have

$$\begin{aligned} \mathbb{E}_{x(h)}[f(x'(\tau))] - f(u, 0, w) &= ah(f(0, z, 0) - f(u, 0, w)) \\ &\quad + \mathbb{E}_{x(h)}[(f(x'(\tau)) - f(u, 0, w)) \mathbb{1}(x'_2(\tau) = 0)] . \end{aligned} \quad (5.8)$$

The first line corresponds to a jump part, where all the mass z ends up in site 2. The second line has to be a continuous part, since the probability of the event $x'_2(\tau) = 0$ is of order $1 + o(h)$. Since $x'_1(t)$ is a martingale we have

$$\mathbb{E}_{x(h)}[x'_1(\tau) - u] = -uah = -uah + \mathbb{E}_{x(h)}[(x'_1(\tau) - u) \mathbb{1}(x'_2(\tau) = 0)] ,$$

where the second equality uses again $\mathbb{P}_{x(h)}[x'_2(\tau) = z] = ah$ and the fact that $x'_2(\tau) = z$ implies $x'_1(\tau) = 0$. Therefore

$$\mathbb{E}_{x(h)}[(x'_1(\tau) - u) \mathbb{1}(x'_2(\tau) = 0)] = 0 \quad (5.9)$$

and similarly for $x'_3(t)$, and the drift of the continuous part vanishes. It remains to compute the diffusion part.

$$\begin{aligned} \mathbb{E}_{x(h)}[(x'_1(\tau) - u)^2] &= \mathbb{E}_{x(h)}[x'_1(\tau)^2] - 2u\mathbb{E}_{x(h)}[x'_1(\tau)] + u^2 \\ &= \mathbb{E}_{x(h)}[x'_1(\tau)(z - x'_2(\tau) - x'_3(\tau))] - 2u^2(1 - ah) + u^2 \\ &= uz(1 - ah) - 0 - uw(1 - ah)^2 - 2u^2(1 - ah) + u^2 \\ &= ahuz + O(h^2) , \end{aligned} \quad (5.10)$$

where we have used that $x'_1(\tau)x'_2(\tau) = 0$ and that $x'_1(t)x'_3(t)$ is a martingale. Again splitting the expectation with respect to the value of $x'_2(\tau)$ we get analogously to the above

$$\mathbb{E}_{x(h)}[(x'_1(\tau) - u)^2 \mathbb{1}(x'_2(\tau) = 0)] = ahuz - ah u^2 = ah u w . \quad (5.11)$$

The same holds for $x'_3(t)$, and obviously $E_{x(h)}[x'_2(\tau)^k \mathbb{1}(x'_2(\tau) = 0)] = 0$ for all $k = 1, 2, \dots$. To get the covariances we compute

$$\begin{aligned} E_{x(h)}[(x'_1(\tau) - u)(x'_3(\tau) - w) \mathbb{1}(x'_2(\tau) = 0)] &= \\ &= E_{x(h)}[x'_1(\tau)x'_3(\tau) \mathbb{1}(x'_2(\tau) = 0)] - uw = -2ahuw, \end{aligned} \quad (5.12)$$

where we used (5.9) and that $x'_1(t)x'_3(t)$ is a martingale. Covariances with $x'_2(t)$ again vanish.

Now, in (5.8) we have up to second order

$$\begin{aligned} f(x'(\tau)) - f(u, 0, w) &= \sum_{i=1}^3 \partial_{x_i} f(x(0))(x_i(\tau) - x_i(0)) \\ &+ \frac{1}{2} \sum_{i,j=1}^3 \partial_{x_i} \partial_{x_j} f(x(0))(x_i(\tau) - x_i(0))(x_j(\tau) - x_j(0)), \end{aligned} \quad (5.13)$$

where higher order terms are of order $o(h)$ and not relevant. Taking expectations, the first order terms vanish, and the second order diffusive terms can be written as (recall that $a = \alpha/4$)

$$E_{x(h)}[(f(x'(\tau)) - f(u, 0, w)) \mathbb{1}(x'_2(\tau) = 0)] = \frac{1}{2} \alpha h u w \partial_{13}^2 f(u, 0, v) + o(h)$$

using the same notation as in (2.5). Plugging this and (5.8) in (5.6) we obtain for the limiting generator A :

$$Af(u, 0, w) = \alpha(f(0, z, 0) - f(u, 0, w)) + \frac{1}{2} \alpha u w \partial_{13}^2 f(u, 0, w), \quad (5.14)$$

consisting of a jump part and a Wright-Fisher diffusive part with effective diffusivity $\alpha u w$. Notice that as soon as u or $w = 0$ the diffusive part vanishes, and the generator consists purely of the jump part where the total mass z moves from site 1 or 3 onto site 2.

5.3 Dynamics of a single condensate: jump process

For the other possible initial condition $x(0) = (0, z, 0)$ we have $x(h) = (zah, z(1 - 2ah), zah)$ as initial condition for the Wright-Fisher diffusion. Again we have

$$\begin{aligned} E_{x(h)}[f(x'(\tau))] &= f(0, z, 0) \mathbb{P}_{x(h)}[x'_2(\tau) = z] \\ &+ E_{x(h)}[f(x'(\tau)) \mathbb{1}(x'_2(\tau) = 0)], \end{aligned} \quad (5.15)$$

where this time the probability in the first line is of order $1 - 2ah$. Since $x'_1(t)x'_3(t)$ is a martingale we have

$$E_{x(h)}[x'_1(\tau)x'_3(\tau)] = (zah)^2 \quad (5.16)$$

which implies that the probability to end up in a state $(u, 0, w)$ with $u, w > 0$ is negligible. Therefore

$$E_{x(h)}[f(x'(\tau)) \mathbb{1}(x'_2(\tau) = 0)] = 2ah \left(\frac{1}{2} f(z, 0, 0) + \frac{1}{2} f(0, 0, z) \right), \quad (5.17)$$

and the generator consists only of jump parts which can be written as

$$Af(0, z, 0) = a(f(z, 0, 0) - f(0, z, 0)) + a(f(0, 0, z) - f(0, z, 0)). \quad (5.18)$$

Therefore, once all the mass concentrates on a single site this remains, and the corner points \mathcal{C} are absorbing for the limit process.

The closure of the operator A defined by (5.18),(5.14) clearly is the generator of a well-defined Markov process on the set \mathcal{A} defined in (5.5). Furthermore, tightness of the sequence of processes generated by L_N (5.2) is proved in Lemma 4.2. Therefore, we can conclude for the three site case, that the processes generated by (5.2) indeed converge to the process generated by the projected generator A defined by (5.18), (5.14). This concludes the proof of Theorems 2.1 and 2.2 for the case of three sites.

6 The general case continued

6.1 Motion of a single condensate

First we show that once a point $e_i \in \mathcal{C}$ is hit in the limit process, i.e., the situation physically corresponding to a single condensate, then the process becomes a pure jump process on the corner set $\mathcal{C} = \{e_i : i \in S\}$. This in combination with Lemma 4.2 finishes the proof of the first item in Theorem 2.1. Recall the decomposition $L_N = L + \theta_N L'$ of the generator of the auxiliary process given in (2.6).

Lemma 6.1. 1. For every Borel measurable subset $K \subset E \setminus \mathcal{C}$ we have for the function $\nu_x(K) : x \mapsto \nu_x(K)$

$$(L\nu_x(K))(e_i) = 0. \quad (6.1)$$

As a consequence, $Af(x) = 0$ for all $x \in E \setminus \mathcal{C}$ and hence,

$$Af(e_i) = \sum_{j \in S} \frac{\alpha}{2} p(i, j) (f(e_j) - f(e_i))$$

i.e., the projected process is a pure jump process on \mathcal{C} .

Proof. Fix $K \subset E \setminus \mathcal{C}$. If $p(i, j) > 0$ then starting from a point $x \in E$ such that $x_k = 0, k \notin \{i, j\}$, the Wright-Fisher diffusion L' remains in the plane $E_{ij} := \{re_i + se_j : r, s \in \mathbb{R}\}$ and hence is absorbed at e_i or e_j . As a consequence

$$\nu_{e_i + se_j - se_i}(K) = 0$$

which implies

$$(\partial_{ij}\nu_x(K))(e_i) = 0.$$

Since K was an arbitrary Borel measurable subset of $E \setminus \mathcal{C}$, we conclude for every bounded Borel measurable function f with support contained in $E \setminus \mathcal{C}$ that

$$\left(\partial_{ij} \int_{\mathcal{A}} \nu_x(da) f(a) \right) (e_i) = 0.$$

Further, if $i \notin \{k, l\}$ then obviously, for any smooth function $f : \mathcal{A} \rightarrow \mathbb{R}$

$$((x_k - x_l) \left(\partial_{kl} \int_{\mathcal{A}} \nu_x(da) f(a) \right) (e_i) = 0,$$

because both x_k and x_l are zero when evaluated at $x = e_i$.

Therefore, since L in (2.9) is a linear combination of terms of the form $-\frac{\alpha}{4}(x_k - x_l)\partial_{kl}$, we conclude that for any smooth function $f : \mathcal{A} \rightarrow \mathbb{R}$ with support contained in $E \setminus \mathcal{C}$

$$L \left(\int_{\mathcal{A}} \nu_x(da) f(a) \right) (e_i) = 0.$$

Therefore, we obtain for all $f \in C(\mathcal{A}, \mathbb{R})$

$$\begin{aligned} Af(e_i) &= L \left(\int_{\mathcal{A}} \nu_x(da) f(a) \right) (e_i) = L \left(\int_{\mathcal{C}} \nu_x(da) f(a) \right) (e_i) = \\ &= \sum_{j \in S} \frac{\alpha}{2} p(i, j) (f(e_j) - f(e_i)), \end{aligned} \quad (6.2)$$

since $\int_C \nu_{e_i}(da)f(a) = f(e_i)$. \square

6.2 The general case of Theorem 2.2

The proof of the general case is a combination of the arguments used in the three site case discussed above, and we will be brief in this section avoiding a full treatment with a lot of technical notation. Recall that $p(i, j) \in \{0, 1\}$ and let $x(0) \in \mathcal{A}$ be a general initial condition, for which $x_i(0)x_j(0) = 0$ whenever $p(i, j) = p(j, i) = 1$. Again, under the Wright-Fisher process $(x'(t) : t \geq 0)$, with generator $L' = (1/2) \sum_{ij} p(i, j) y_i y_j \partial_{ij}^2$, all $x'_i(t)$ and all products $x'_i(t)x'_j(t)$ for which $p(i, j) = 0$ are martingales. Analogously to the computations in Section 5, after the action of the deterministic drift generator $L = \sum_{ij} -ap(i, j)(y_i - y_j)\partial_{ij}$ for infinitesimal time h we have an initial condition $x(h)$ for the diffusion L' , and denote by τ the time of absorption.

The first observation is that at most one initially empty site can gain a macroscopic amount of mass through a jump.

Lemma 6.2. *$x_i(0) = 0$ and $x'_i(\tau) > 0$ is possible for at most one $i \in S$, and in that case $\sum_{j \in S} p(i, j)x'_j(\tau) = 0$.*

Proof. Let i and j be two initially empty sites for which $x_i(h), x_j(h) = O(h)$. Then, if $p(i, j) = 1$ we have $x'_i(\tau)x'_j(\tau) = 0$ since τ is the hitting time of the absorbing set \mathcal{A} , so that at most one of them can take a positive value. Similarly, if $p(i, j) = 0$ we have $x'_i(\tau)x'_j(\tau) = O(h^2)$ since the product is a martingale, and since we take the limit $h \rightarrow 0$ again at most one can be positive after the diffusion is absorbed. Since the pair i, j was arbitrary, this implies that under the action of the generator A at most one empty site in the system can gain mass. Then at the time of absorption in \mathcal{A} all other connected sites have to be empty. \square

Secondly, jumps of a macroscopic amount of mass occur only onto neighbouring empty sites.

Lemma 6.3. *If $x_i(0) = 0$ and $\sum_{j \in S} p(i, j)x_j(0) = 0$, then $x'_i(\tau) = 0$. If $x_i(0) > 0$ then $x'_i(\tau) \leq x_i(0) + O(h)$.*

Proof. By the condition site i has no direct neighbour with $x_j(0) > 0$, and therefore if $x_i(0) = 0$ under the action of generator L in time h we get $x_i(h) = O(h^2)$. Then since $x'_i(t)$ is a martingale, its mass remains negligible after the Wright-Fisher diffusion L' gets absorbed and $x'_i(\tau) = x_i(0) = 0$.

If $x_i(0) > 0$ we also have $\sum_{j \in S} p(i, j)x_j(0) = 0$ and thus $x_i(h) \leq x_i(0) + O(h^2)$. Then the martingale property of x'_i implies that absorption at $x'_i(\tau) = x_i(0) + O(1)$ has negligible probability of order h^2 . \square

Thirdly, mass between two macroscopically occupied sites moves at most continuously.

Lemma 6.4. *If $x_i(0)x_j(0) > 0$, then either $x'_i(\tau)x'_j(\tau) = 0$ or $x'_i(\tau)x'_j(\tau) = x_i(0)x_j(0) + O(h)$.*

Proof. $x_i(0)x_j(0) > 0$ implies $p(i, j) = 0$. In the first case the mass of one of the sites jumped to another site. Now assume this is not the case and note that $x'_i(t)$, $x'_j(t)$ and $x'_i(t)x'_j(t)$ are martingales. Let $f, g : [0, 1] \rightarrow \mathbb{R}$ be two smooth functions, then

$$L'(f(x_i)g(x_j)) = \sum_{k \in S} \left(p(i, k)x_i x_k \partial_{ik}^2 f(x_i) + p(j, k)x_j x_k \partial_{jk}^2 g(x_j) \right). \quad (6.3)$$

Then $x_i(h)x_j(h) = O(1)$ which implies that $x_k(h) = O(h)$ for all k connected to either i or j and

$$L'(f(x_i(h))g(x_j(h))) = O(h). \quad (6.4)$$

So either a neighbour $x'_k(t)$ gains a macroscopic amount of mass during the Wright-Fisher diffusion, which results in the first case of one of the piles jumping to k , or we have

$$\mathbb{E}_{x(h)}[f(x'_i(\tau))g(x'_j(\tau))] = f(x_i(0))g(x_j(0)) + O(h). \quad (6.5)$$

Since f and g are arbitrary, the latter case implies the second statement. \square

Lemma 6.5. *If $x_i(0) = 0$ and $x'_i(\tau) > 0$ then $x'_i(\tau) = \sum_{j \in S} p(j, i)x_j(0) + O(h)$.*

Proof. This follows from Lemmas 6.2 and 6.4. i is the only site in the system where a macroscopic amount of mass jumped to and its neighbours are empty at absorption, therefore it has to absorb all the original mass of its neighbours up to amounts of order h which have been shared with other sites. \square

This characterizes all the possible jump events under the generator A . We see that mass can only jump onto a single empty site j which collects all the surrounding mass, and the rate for this event can be computed analogously to the previous section to be $\frac{\alpha}{2} \sum_{i \in S} p(i, j)x_i(0)$. In parallel to jumps, mass can move continuously between macroscopically occupied sites, which we describe next. We will use the notation $\hat{p}(i, j)$ introduced in (2.15) for the number of two-step connections between two sites $i, j \in S$. The crucial ingredient to conclude is additivity of the jump rates of the Wright-Fisher diffusion in each component, as is described below.

Lemma 6.6. *Assume that $\hat{p}(i, j)x_i(0)x_j(0) > 0$ then the generator A contains a term $a\hat{p}(i, j)x_i(0)x_j(0)\partial_{ij}^2$ and the continuous part of A is given by the sum of all such contributions.*

Proof. For $i, j \in S$ as given, assume that $x'_i(\tau)x'_j(\tau) > 0$ so that none of the two piles jumps and the total mass is conserved up to order h , which has $O(1)$ probability. Then the system on sites i, k, j with intermediate site k can be seen as an effective three site system as studied in the previous section. If there is a unique intermediate site, the diffusive part of the generator follows analogously to this case. If there are several intermediate sites k , the different paths between i and j can be viewed as independent, since the diffusion rates of the Wright-Fisher generator acting on sites i and k given by

$$L'_{ik} = x_i x_k \partial_{ik}^2 \quad (6.6)$$

are proportional to x_k (similarly between j and k). Therefore the total effective Wright-Fisher diffusion rate between sites i and j is simply given by a sum over all connections, as given by $\hat{p}(i, j)$ in (2.15). Since the above rates are also linear in x_i (and x_j , resp.), the same additivity applies if i or j are connected to more different sites via two-step connections. This leads to independent contributions for each such connection as claimed in the Lemma, which are summed over to give the full continuous part of the generator.

It remains to show that there is no mass exchange over distances more than 2 steps. This is most easily seen in a system with 4 sites $S = \{1, 2, 3, 4\}$ with nearest neighbour jumps and initial condition $(u, 0, 0, w)$ and $u + w = z$. Analogous to the computation in (5.10) for three sites we get

$$\mathbb{E}_{x(h)}[(x'_1(\tau) - u)^2] = ah(uz - uw) + O(h^2).$$

This implies analogously to (5.12) that

$$\mathbb{E}_{x(h)}[(x'_1(\tau) - u)^2 \mathbb{1}(x'_2(\tau) = 0)] = ah(uz - uw) - ah u^2 + O(h^2) = O(h^2),$$

so the change of x_1 is not on a diffusive scale and negligible. The same holds for x_4 , and this argument can be easily generalized to arbitrary S . \square

References

- [1] Giardinà, C., Kurchan, J. and Redig, F.: Duality and exact correlations for a model of heat conduction. *J. Math. Phys.* **48**, (2007), 033301. MR-2314497
- [2] Giardinà, C., Redig, F. and Vafayi, K.: Correlation inequalities for interacting particle systems with duality. *J. Stat. Phys.* **141**, (2010), 242-263. MR-2726642
- [3] Grosskinsky, S., Redig, F. and Vafayi, K.: Condensation in the inclusion process and related models. *J. Stat. Phys.* **142**(5), (2011), 952-974. MR-2781716
- [4] Moran, P.A.P.: The statistical processes of evolutionary theory. *Clarendon Press*, Oxford, 1962.
- [5] Arnoldt, H., Timme, M. and Grosskinsky, S.: Frequency-dependent fitness induces multistability in coevolutionary dynamics. *J. R. Soc. Interface* **9**(77), (2012), 3387-3396. arXiv:1209.2638
- [6] Giardinà, C., Kurchan, J. and Redig, F. and Vafayi, K.: Duality and hidden symmetries in interacting particle systems. *J. Stat. Phys.* **135**, (2009), 25-55. MR-2505724
- [7] Bernardin, C.: Superdiffusivity of asymmetric energy model in dimensions 1 and 2. *J. Math. Phys.* **49**, (2008), 103301. MR-2464620
- [8] Waclaw, B. and Evans, M.R.: Explosive condensation in a mass transport model. *Phys. Rev. Lett.* **108**, (2012), 070601. arXiv:1110.4330
- [9] Hirschberg, O., Mukamel, D. and Schütz, G.M.: Motion of condensates in non-Markovian zero-range dynamics. *J. Stat. Mech.*, (2012), P08014. arXiv:1207.1006
- [10] Jara, M. and Beltrán, J.: private communication
- [11] Beltrán, J. and Landim, C.: Metastability of reversible condensed zero range processes on a finite set. *Probab. Theory Relat. Fields* **152**(3-4), (2012), 781-807. MR-2892962
- [12] Landim, C.: Metastability for a non-reversible dynamics: the evolution of the condensate in totally asymmetric zero range processes, arXiv:1204.5987
- [13] Pavliotis, G.A. and Stuart, A.: Multiscale Methods - Averaging and Homogenization. *Springer*, Berlin, 2008.
- [14] Kurtz, T.G.: A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Funct. Anal.* **12**, (1973), 55-67. MR-0365224
- [15] Grosskinsky, S., Schütz, G.M. and Spohn, H.: Condensation in the zero range process: stationary and dynamical properties. *J. Stat. Phys.* **113**(3/4), (2003), 389-410. MR-2013129
- [16] Godrèche, C. and Luck, J.M.: Dynamics of the condensate in zero-range processes. *J. Phys. A* **38**, (2005), 7215-7237. MR-2165701
- [17] Liggett, T.M.: Interacting Particle Systems. *Springer*, Berlin, 1985. MR-2108619
- [18] Huillet, T.: A Duality Approach to the Genealogies of Discrete Nonneutral Wright-Fisher Models. *J. Probab. Stat.*, (2009), 714701. MR-2602875
- [19] Whitt, W.: Proofs of the martingale FCLT. *Prob. Surveys* **4**, (2007), 268-302. MR-2368952
- [20] Ethier, S.N. and Kurtz, T.G.: Markov Processes - Characterization and Convergence. *Wiley*, 1986. MR-1567615
- [21] Huillet, T.: On the Karlin-Kimura approaches to the Wright-Fisher diffusion with fluctuating selection. *J. Stat. Mech.*, (2011), P02016.

Acknowledgments. S.G. acknowledges support by the Engineering and Physical Sciences Research Council (EPSRC), Grant No. EP/I014799/1.